

Interactive out-of-core visualization of very large multiresolution time series data

PI: R Daniel Bergeron, U of New Hampshire

One of the major challenges facing scientists today is effective analysis of the explosion of data that is being generated by current scientific research technology. This is particularly a problem in the simulation of time-dependent magnetohydrodynamics (mhd) phenomena that produces enormous data sets of many hundreds of gigabytes. It is also the case that interactive visualization of the data is still one of the most effective techniques for gaining insight into the physical phenomenon. The huge size of the data, however, severely limits the ability to visualize the data in an effective interactive environment. A promising solution to this dilemma is to represent the data at multiple resolution levels - the coarser resolutions are used for large-scale overview visualizations and increasingly finer resolutions are accessed as the user zooms in for more limited range views. This project builds on previous work by the computer science principal investigators in the areas of multiresolution representation of scientific data and multidimensional caching and prefetching. We propose to develop a software framework and a prototype visualization environment that will support the interactive visualization of three-dimensional time-dependent mhd simulations using multiresolution and adaptive resolution data representations. An integral component of this research will be the use of very efficient I/O caching and prefetching techniques that are critical for achieving interactive performance. Although the conceptual framework applies to a very wide range of scientific data, the prototype visualization environment will focus on space science applications, particularly magnetohydrodynamics simulation and computational fluid dynamics. These areas represent two extremely important science research domains within the NASA research mission. The proposed work specifically addresses the AISR goal of providing better tools to enhance the scientists' ability to do research.

Extending Virtual Observatory Capabilities Through Grid Enabled Resources

PI: Adam Szabo, NASA Goddard

Existing and forthcoming Virtual Observatories will enable unified access to many disparate data sources and services. As these systems expand they face issues in scalability. In particular, user contributed services and applications, which play a vital role in long-term success, are hindered by volume of use and lack of hosting facilities. We aim to address this issue by establishing the infrastructure for a VO-Grid. The Globus Toolkit, the de facto open source standard in Grid software, will be used to establish a mechanism for allowing user contributed submissions of services. Additionally, the infrastructure for allowing these services to be shared along with unused processor time and storage space will be set in place.

Global cyclone detection and tracking (GLYDER) for climate variability studies: A multisensor data processing approach

PI: Ashit Talukder, NASA JPL

We will implement an automated data processing and pattern recognition system, termed GLYDER (GLObal cYclone DETection and tRacking), that classifies and tracks cyclones in multi-sensor satellite data sets. This system will be used to quantify the spatial and temporal variability of historical cyclones and their tracks globally and thereby supports a key goal of NASA's Earth Science Enterprise (ESE) to quantify the Earth's natural variability and understand how it is changing. Cyclone detection with high accuracy using only single data modalities is an unsolvable problem, largely due to variations in the types of cyclones, and the presence of other weather patterns that appear similar within the sensor measurements. Incorporating multiple sensor types together provides several cues that can better characterize cyclones. The joint temporal-spatial structure of cyclones has not been exploited in previous efforts and we believe that using such constraints will significantly improve detection accuracy over NCEP analysis, while reducing false alarms. In GLYDER, we will demonstrate the first multimodal cyclone detection scheme from multiple remote sensors with varying spatio-temporal resolution. We will develop and test a software system that will co-register multiple remote sensory data (GOES, QuickScat, AVHRR, SSMI and AMSR-E) that are at different spatial and temporal resolutions and detect and track cyclones from co-registered multiple remote sensors using advanced proven pattern recognition and tracking capabilities. Our goal in GLYDER is to co-register & detect cyclones at a spatial resolution of 10 kms and a temporal resolution of 3 hrs. We will use advanced multiscale wavelet techniques to co-register disparate multiple satellite datasources at different temporal and spatial resolutions. We will then use proven pattern recognition and data mining Maximum Discriminant Function Classifiers (MDF) combined with robust feature tracking algorithms on the remote multisensor co-registered data to detect cyclones and then track them temporally as they evolve over time. Based on previous performance of our detection and tracking algorithms on other datasets, we estimate that GLYDER will enable at least a 50% improvement in automated cyclone detection/classification from remote sensors with a 4X reduction in false alarms compared to current NCEP-based single sensor cyclone detection. We will initially train, test and validate our algorithms on regions that have multiple in-situ measurements and also experience frequent and intense storms and hurricanes (tropical eastern Atlantic and northeast pacific). Finally we will test and demonstrate our system with data from the Southern Oceans. The GLYDER system will empower scientists with the capability to detect global cyclones in long time series of data. GLYDER also has the potential for use in an operational environment for early detection of tropical storms. Another arena in which GLYDER can play an important role is that of knowledge search and discovery. Current NASA's GCMD satellite data search support only simple searches for satellite measurements in time & space only. To facilitate advanced content-based searches, the NASA PO.DAAC would use GLYDER to mine its data holdings for the global cyclones and then tag the metadata of each granule with properties related to the cyclone for use in systems such as ECHO. As part of the NASA AISR program, GLYDER will be utilized by: 1) JPL climate scientists to study and quantify the spatial and temporal variability of cyclones and their tracks (a critical need that has been identified by the Intergovernmental Panel on Climate Change); 2) NASA's Physical Oceanography Distributed Active Archive Center (PO.DAAC) to tag metadata with information pertaining to cyclones and enable content-based searching.

Improving Remote Sensed Data Products Using Bayesian Methodology for the Analysis of Computer Model Output

PI: Robin Morris, USRA-RIACS

The NASA Science Directorate supports an array of Earth Observing satellites which provide global coverage, and whose observations are used to produce a wide array of data products, from measurements of sea surface temperature, to polar ice coverage, to plant type and growth rates. These data products are used in a wide variety of further scientific studies, and also as inputs to important policy decisions, especially those concerning the impact of human activity on the biosphere. Many of the studies of human impact concern changes over time. To accurately characterize changes it is vitally important that the uncertainties in the estimates of the quantities being observed are known, so that the uncertainty in the estimated changes can be accurately determined -- making scientific or policy decisions based on estimates with large, or worse, unknown or poorly determined errors, is poor science and poor policy. Many of the data products (e.g. Leaf Area Index, Net Primary Production, Photosynthetically Active Radiation from MODIS (Moderate Resolution Imaging Spectroradiometer)) are produced by inverting a Radiative Transfer Model (RTM), which simulates the upwelling radiation at the top of the atmosphere (and so observed by the satellite) as a function of the biospheric parameters (e.g. land cover type; available water; leaf chemistry; etc.). These RTMs are implemented as complex computer codes, and the analysis and inversion of these codes is a challenging task. In the last several years the area of Bayesian Methods for the Analysis of Computer Model Output has made great progress, and is coming to a point where its wider application will show significant utility in the application domains. These methods are well developed in the statistics literature, but are almost unknown in the Geoscience/Remote Sensing domain (with the exception of Kriging). Applied to an RTM, these methods will allow the determination of: a) the uncertainty in the RTM output; b) the main effects, that is, which of the inputs is mainly responsible for the output uncertainty; c) validation using field data; d) rapid approximation of the RTM for use when computing the inverse; e) a direct model for the inverse incorporating uncertainty. Advances in these areas will improve the accuracy and utility of the data products. The RTM in operational use for estimating LAI (leaf area index) and fPAR (fraction of photosynthetically active radiation) is the MOD15 algorithm. We propose initially to apply the methodology to the LCM (leaf-canopy model) RTM, a similar model to which we have more direct access. The methods and code developed will be readily applicable to MOD15 and other RTMs -- the RTM is treated as a "black box" by the methodology. Initially, a Gaussian Process (GP) model will be built using runs of the LCM. Sensitivity analysis using the GP model approximation will be performed to determine where more information is required about the distributions of the input variables. The fast computations available with the GP model will enable practical inversion of the LCM, to determine the distribution over LAI given measured reflectances. Computing the distribution over LAI for each observation using the forward model as described above may not be sufficiently computationally cheap to implement over large areas. We will build a second GP model that directly models the inverse problem, using training data generated from the full inverse. This will model the relationship between satellite observations and the probability distribution over LAI, and will demonstrate a new generic approach to characterizing the uncertainty in inverse problems. A successful conclusion of the work in this proposal would demonstrate the utility of the methodology in an important application domain, with significant scientific and policymaking consequences, and would significantly advance the state-of-the-practice in remote sensing science.

Building and Implementing a Virtual Sun-Solar System (Great) Observatory (VS3O)

PI: Aaron Roberts, NASA Goddard

We propose to build on the success of the Virtual Space Physics Observatory (VSPO) and related efforts efforts to realize the potential of Virtual Observatories to unite data and service providers in the Sun-Solar System domains into a new "Great Observatory," VS3O. Current plans at NASA and in the international, multi-agency Sun-Solar System Connection community include making data easily available from all missions relevant to the global problem of the effects of solar plasma dynamics on the Earth and other planets. To accomplish this, Virtual Observatories register data products and services from disparate providers using a common language that allows searching and use in a uniform way. Subfields may be organized along the lines of the now operational Virtual Solar Observatory, but more is needed to realize the full potential of these systems. We will solve the following critical problems in implementing VS3O: (1) developing a suite of software to allow resource providers to easily transform whatever system they have for accessing their data into a service for VOs; (2) providing an API and useful visual and other (e.g., IDL) front-ends to the system; (3) building on current efforts to provide a "mission independent data layer" to allow wide ranges of data and models to be dealt with uniformly; (4) integrating the independently developed "higher order query" service along with subsetting and other other services we will develop into VS3O; (5) incorporating parallelism as needed to make the VS3O responsive to large numbers of complex queries; (6) developing community outreach, feedback, and oversight mechanisms that assure broad participation and high utility; and (7) beginning efforts to incorporate the resulting Observatory into public and educational outreach. This system will be flexible, flexible, extensible by others, and will greatly aid in doing focused research projects as well as facilitating large-scale research efforts. We will develop operational modules for various functions that will also serve as prototypes for others to use the VS3O in currently unforeseen ways. The PI and team we have assembled has strong connections to most of the current efforts in this area and has a proven track record of fundamental accomplishments.

Auroral Phenomenon Localization, Classification, and Temporal Evolution Tracking

PI: Timothy Newman, UA Huntsville

In this project, techniques that will aid investigation of auroral phenomena will be developed. The techniques to be developed will be automated and shape model-based and will allow accurate localization of the auroral oval (in the presence of unfriendly background noise) from UVI and IMAGE FUV data. The techniques will provide useful input to downstream analysis tasks, including classification of auroral features and exploration of auroral feature temporal evolution. New, robust automated techniques for feature classification and for exploration of temporal evolution will also be developed. All of these techniques will be incorporated into a unified auroral phenomenon search tool.

Innovative Techniques for Producing Corrected Synoptic Maps for a UVCS Outflow Velocity Database

PI: Leonard Strachan, Harvard/SAO

The primary objective of this project is to develop a set of software tools that will allow researchers to obtain empirically derived plasma parameters of the solar wind acceleration region of the corona. We will produce line-of-sight corrected synoptic maps of outflow speeds and kinetic temperatures for both hydrogen and O5+ ions in the extended solar corona. Such information is valuable for constraining theoretical solar wind models and it can be used in Space Weather models that predict ambient solar wind conditions. A set of visualization tools for displaying the derived coronal parameters will also be developed and made available to the public. The derived outflow velocities will be based on well-proven UVCS/SOHO spectroscopic techniques. Synoptic maps of various parameters will be produced using software that has been developed for solar tomography applications. Extensive use will be made of open source software for developing visualization tools for displaying the data. All software developed from the project will be applicable to other NASA programs that require the use of 3-D data rendering such as the STEREO mission.

Monitoring and Diagnosis of Complex Software and Hardware for Earth Observing Missions

PI: Brian Williams, MIT

The Autonomous Sciencecraft Experiment (ASE) onboard the Earth Observing One (EO-1) mission has demonstrated the potential of autonomous systems to maximize scientific return. As complex software systems are being developed to control space assets and optimize onboard operations, there is a growing need for verification and validation of these systems. However, verification is traditionally performed offline during design and development, and does not guarantee a safeguard from all possible system failures. To complement offline verification techniques and ensure extremely high reliability operations, we propose to develop an onboard model-based fault management system to monitor and diagnose complex software and hardware systems, and track the progress of high-level mission objectives, in the context of the EO-1 mission. Our proposed technology will extend previous hardware diagnosis engines, engines, such as Livingstone, to mixed hardware and software systems in several ways. First, it will monitor embedded software and diagnose software anomalies to enable robust execution and maximum science return. Second, monitoring software state will be used for refining the diagnosis of hardware components. Finally, the proposed engine will be capable of diagnosing in the presence of delayed symptoms, for the general case of mixed hardware and software systems. We will demonstrate the capability of the new model-based fault management by diagnosing ASE software and the progress of the high-level scientific goals. To assure the authenticity of the demonstration, we will integrate this fault management system with the ASE software and test by simulating the actual software anomalies detected during the execution of ASE onboard the EO-1 mission. We additionally propose an optional 12 month extension to flight validate the capability on EO-1. This proposal directly responds to the Applied Information Systems Research (AISR) program objectives of NASA's Research Opportunities in Space and Earth Sciences (ROSES). In particular, our technology is expected to enhance the science productivity of NASA's space flight missions that are sponsored by the Science Mission Directorate (SMD). The proposal will build upon the success of the ASE onboard the EO-1 mission, by providing an onboard capability for monitoring and diagnosing software and hardware systems, based on lessons learned from the ASE. Enhancing the ASE software through the proposed fault management capability will enable extremely high reliability operations, resulting in an increased return of scientific data. This proposal also responds to NASA's Strategic National Objective to $\frac{1}{2}$ Study the Earth system from space and develop new space-based and related capabilities for this purpose. $\frac{1}{2}$ The maturation and validation of our proposed technology in the context of EO-1 will demonstrate its potential for long term impact on many future NASA missions that are increasingly relying on complex software and hardware systems.

Novel Higher-Order Statistical Method for Extracting Dependencies in Multivariate Geospace Data Sets

PI: Jay Johnson, Princeton U

Understanding magnetospheric dynamics and the relationship between the solar wind driver and the magnetospheric response is of great practical interest because it could potentially help to avert catastrophic loss of power and communications. In order to build good predictive models it is necessary to understand the most critical nonlinear dependencies among observed plasma and electromagnetic field variables in the coupled solar wind/magnetosphere system. We propose to develop and apply a novel cumulant-based information-dynamical measure in conjunction with other nonlinear techniques to magnetospheric data to characterize: (1) the underlying dynamics, (2) the nonlinear behavior of the geospace systems, (3) the solar cycle dependence of these properties, and (4) the predictability of the systems. Because this nonparametric, statistical approach assumes no intrinsic underlying dynamics, it is possible to avoid some of the pitfalls involved in parametric modeling and/or physics based models. Moreover, the information gained from the cumulant-based information flow could be used to guide the development of predictive models. To illustrate the power of our method, we will examine the nonlinear cross dependencies in a large database of geospace variables. We will evaluate the underlying dynamics of the magnetosphere by examining the time evolution of geomagnetic indices, Kp, Dst, and AE, as well as more direct measures of the magnetospheric state, energetic electron flux, tail stretching, and energetic electron precipitation. From this data we will be able to estimate a predictability horizon which will indicate the maximum "look ahead" for which the space climate can be predicted. We will also use the information-dynamical measures to identify variables derived from solar wind data that maximize the information content about the magnetospheric response. This work should have a significant impact on space science, space/mathematical science education, and society at large. First, we have developed a new technique for analyzing nonlinear dependencies in large data set using cumulant-based significance measures. Because the methods and techniques do not presuppose an underlying dynamics, they are transportable across a wide range of complex nonlinear systems which can broadly impact fields ranging from geoscience to bioscience. Second, detection of the most important nonlinear interactions in the systems could be used to identify the most important physical processes that will aid development of physical models and predictive capabilities for understanding severe space weather. Third, we will submit an E/PO proposal where we will develop web-based outreach tools to teach high school students basic concepts in nonlinear dynamics with applications to space science. Finally, the results will be widely disseminated in mathematical and geospace journals and presented at meetings and workshops.

Sun Earth Connection Distributed Data Services

PI: Peter Cornillon, OPeNDAP

The goal of the work proposed herein is to integrate a number of current IT efforts into a suite of data services designed to facilitate the access to and use of data for Sun Earth Connection studies. The collection of services to be developed is based on the OPeNDAP Data Access Protocol (DAP), a protocol in heavy use in the Earth sciences. Although these services will be developed for Sun Earth Connection studies, we anticipate that they will be of use to other areas in the sciences as well. Specifically, we will make use of some of these enhancements in ongoing data system work in oceanography; i.e., this project will leverage work done in the Earth sciences to benefit space physics and the work done to benefit space physics will feed back to the Earth sciences. General Web services to be developed or enhanced as part of this effort: - High bandwidth data access - OPeNDAP-g, a Grid-enabled version of OPeNDAP, will be re-integrated into the OPeNDAP core such that all OPeNDAP core services are Grid accessible. - Catalog access integrated with content-based access - A hierarchical cataloging capability being developed at UCAR will be integrated into the OPeNDAP core, greatly facilitating access to large multi-file data sets. - Metadata services - The ability to augment data set descriptions at a central site for remotely accessible data sets will be enhanced to allow for the addition of variables and for data object type modification. This will make it possible to provide consistent, 'complete' descriptions of data sets of interest without requiring that the metadata at the data provider's site be modified. Specialized capabilities to be developed for Sun Earth Connections science: - Metadata - Data set descriptions will be completed for Sun Earth Connections data sets of broad interest to the community and made accessible via an OPeNDAP Ancillary Information Services server at HAO. - Data servers - An SPDML-OPeNDAP server will be written. This will provide a base for the metadata descriptions to be used in the project while at the same time greatly facilitating the provision of project related data via OPeNDAP. - Application clients - The IDL-OPeNDAP client will be rewritten to work on a broader range of CPUs and specialized IDL scripts to further facilitate access and use of key data sets will be developed. - Data server installation - OPeNDAP data servers will be installed at several sites for a targeted suite of Sun Earth Connections data sets.

Machine Learning and Data Mining for Automatic Detection and Interpretation of Solar Events.

PI: Arthur Poland, George Mason University

We propose to use novel information technology (IT) methods, characteristic of machine learning and data mining, to address issues related to space science data. In particular, we will develop the capability to automatically detect and characterize solar events for purposes related to space weather. The detection and tracking of solar events is important for the Sun-Earth Connection study and space weather prediction. As the amount of data available becomes overwhelming and near-real time detection is important, there is a strong need to automate this process. The specific tasks addressed by this proposal are to use techniques already developed for other problems to automate the detection, tracking and characterization of solar events and to uncover meaningful relationships between different solar events. Towards that end we propose to develop automated and interdisciplinary computational means driven by machine learning, data mining, image processing, pattern recognition and computer vision. The main modules for our automated system include: (A) preprocessing to improve the quality of data using morphology and relaxation tools, (B) detection of solar objects using a combination of methods including boundary detection, region growing and wavelet analysis, (C) tracking using Condensation (Conditional Density Propagation) methods, (D) characterization of solar event, and (E) spatial/temporal data mining of associations between different type of solar events. The proposed methodology is iterative and the steps (modules) reinforce each other in their findings. We will develop the software tools needed and test the proposed system using, among others, CME observations from LASCO (Large Angle Spectrometer and Coronagraph) and coronal observations from EIT (Extreme-ultraviolet Imaging Telescope) instruments on the SOHO spacecraft. This proposal investigates novel information technology and computational methods to increase productivity of the OSS missions and programs. It will be achieved through close collaboration between experts in computational sciences, computer science and space science.

Robust Localization, Classification, and Temporal Evolution Tracking in Auroral Data

PI: Tim Newman, University of Alabama, Huntsville

A project to develop techniques that aid investigation of auroral phenomena is proposed. The project involves exploiting shape-based techniques that can enable accurate localization of the auroral oval (in the presence of unfriendly background noise), classification of auroral features, and auroral temporal evolution tracking from UVI and IMAGE FUV data, and wrapping these techniques into a web-based content-based image retrieval system. The work includes development of new, robust automated techniques for feature classification and for exploration of temporal evolution.

Solar Loop Mining to Support Studies of the Coronal Heating Problem

PI: Olfa Nasraoui, University of Louisville

The Coronal Heating Problem is the longest standing unsolved mystery in astrophysics. Measurements of the temperature distribution along the loop length can be used to support or eliminate various classes of coronal temperature models. The temperature analysis of coronal loops is a state-of-the-art astronomy. In order to make progress, scientific analysis requires data observed by instruments such as EIT, TRACE, and SXT. The combination of EIT, TRACE, and SXT information provides a powerful data set that will yield unprecedented detail on the plasma parameters of a variety of coronal loop structures. The biggest obstacle to completing this project is putting the data set together. EIT has taken over 300,000 images during its 6-year (and counting) mission. The search for interesting images (with coronal loops) is by far the most time consuming aspect of this project. Currently, this process is performed manually, and is therefore extremely tedious, and hinders the progress of science in this field. The next generation "EIT" called MAGRITE, scheduled for launch in a few years on NASA's Solar Dynamics Observatory, should be able to take 300,000 images in about four days! and will no doubt need state of the art techniques to sift through the massive data to support scientific discoveries. We propose an approach based on data mining to quickly sift through massive data sets downloaded from the online NASA solar image databases and automatically discover the rare but interesting images with solar loops, which are essential in studies of the Coronal Heating Problem. The proposed solar loop mining scheme will rely on the following components: 1) Collection and labeling of a sample data set of images coming from both categories (with and without solar loops), 2) An optimal feature selection strategy that will facilitate the retrieval task, 3) A classification strategy to classify the transformed image into the correct class, and 4) Appropriate measures to validate the effectiveness of the loop mining process. This project will be implemented in three main phases that target the image databases collected by two different instruments, EIT aboard the NASA/European Space Agency spacecraft SOHO and NASA's TRACE. We will leave open the possibility of targeting the SXT database on the Japanese Yohkoh spacecraft if time permits. Each phase will involve collecting and labeling a small sample of the images from the targeted instrument, for which a solar loop mining strategy is to be designed. All the results of this project: literature, software, and outputs Semantic loop features and class labels (in `<emph{XML}</emph>` format) of the developed classification methods on tested portions of the different instrument databases will be made available to the public and other interested researchers via the World Wide Web.

Advanced MHD Algorithm for Solar and Space Science

PI: Dalton D. Schnack, SAIC

The Living With a Star initiative aims to understand those aspects of the Sun-Earth system that affect our life and society. It will investigate the influence of magnetic reconnection on the global corona, the initiation and propagation of CMEs, the evolution of active regions, and the identification of the birthplace of transients. It promises to deliver images and magnetic field measurements with an unprecedented resolution. New advanced software algorithms are required to model these observations and assimilate observational data into models. The magnetohydrodynamic (MHD) model is appropriate for describing the low-frequency motions of the coronal plasma. In order to be useful to investigate the science problems in the LWS program,, an MHD algorithm must satisfy several computational requirements due to: *Extreme separation of time scales. *Extreme separation of spatial scales. For the algorithm to be widely adopted by the solar and space science community, it must: *Use parallel computing methods. *Use standard languages and software. We propose to meet these requirements through an open source project to deliver a sophisticated MHD algorithm highly reliant on other free and/or open source software tools. The goal of the project is to develop MH4D: parallel, machine independent, code that solves the resistive MHS equations on an unstructured grid of tetrahedra. It will employ Adaptive Mesh Refinement (AMR) to resolve evolving small scale spatial structures, such as current sheets and shock waves. It will feature a fully implicit treatment of viscous and resistive dissipation, and a semi-implicit advancement of the momentum equation (the time step will be limited only by the advective CFL condition). MH4D will be developed in Fortran 90 and MPI, use publicly available libraries and software tools such as PETSc, Metis, and GMV, and be made accessible to the solar and space science community.

Inteball and Vega Data Archiving

PI: David Sibeck, JHU/Applied Physics Laboratory

We propose to complete our efforts to calibrate and combine interplanetary, magnetosheath, and magnetospheric observations from the two Interball-1 magnetometers. We will serve the merged data sets from mirrored sites at our home institutions (IZMIRAN and JHU/APL) and work with NASA/NSSDC personnel to add them to the interactive CDAWeb server. The resulting data set will be an integral part of NASA's ISTP mission. In addition, we propose to translate, calibrate, and despoke the Vega-1 and -2 interplanetary magnetometer observations. We will establish a WWW server and work with NSSDC personnel to add the Vega observations to the interactive COHOWeb data server, thereby completing the archival of all relevant interplanetary observations from these heliospheric spacecraft.

The Analysis and Archival of HRDI data in the MLT region

PI: Wilbert R. Skinner, University of Michigan Ann Arbor

The High Resolution Doppler Imager (HDRI) on the Upper Atmosphere Research Satellite (UARS) has measured the wind, temperature, and ozone density in the middle atmosphere regions for approximately 10 years. The HRDI data, along with the rest of the UARS data set, comprise a unique record of the middle atmosphere. As the UARS project winds down it is appropriate to pursue measures to preserve and make this data available to future researchers in ways that are most useful. We propose to enhance the value of this data set by: 1) Converting all levels of HRDI data to a common format (netCDF) and making them available to the scientific community over the internet and on DVD; and 2) incorporating the Tide-Mean Assimilation Technique (TMAT) model with the HRDI data to obtain meaningful estimates of the tides and mean conditions over short time periods. This data set in a convenient form will be very useful for investigators studying the mesosphere and lower thermosphere *MLT) and particularly for those working on the TIMED data set.

Terrestrial Plasmasphere Feature Tracking and Volume Visualization via Tomographic Backprojection

PI: Tim Newman, University of Alabama, Huntsville

The proposed work will enable more accurate and useful exploitation of satellite data of the terrestrial plasmasphere. A recently-developed tomographic backprojection technique that allows reconstruction of plasma densities in 3-space from a single satellite image will be enhanced by coupling with conservative models of plasma distribution in the ionosphere and inner plasmasphere. Integration of data from multiple viewpoints will also be utilized to produce derived volume data products suitable for scientific study. In addition to allowing the volumetric visualization of plasma distributions in the plasmasphere by traditional volume rendering methods, the work will enable plasma features to be manually observed. To aid the observation process, methods to track plasma features in the reconstructed images and volumes will also be developed.

Bibliography

Snake-based Technique for Plasmapause Tracking Kandimalla, R.; T. Newman; Proc., Int'l Conf. on Pat. Rec., Cambridge, UK, (August 2004)

Exploiting Known Latitudinal Variations for Improved Limited-Angle Tomographic Reconstruction of the Plasmasphere Santhanam, N.; T. Newman; D. Gallagher; Proc., Int'l Conf. on Info Tech (ITCC '04), Las Vegas, (April 2004) pp. 602-606

New Hybrid Technique for Volume Reconstruction from Limited Views through an Optically-thin Medium Zhang, H.; T. Newman; Proc., Int'l Conf. on Imaging Science, Systems and Technology (CISST '03) Las Vegas, (June 2003) pp. 277-282

Next Generation Transport Phenomenology Modeling Tool for Terrestrial and Planetary Applications

PI: Douglas J. Strickland, Computational Physics Incorporated

In this project, previously developed and validated codes for modeling photoelectron and auroral electron transport and related optical emissions for the upper atmosphere are being reengineered and incorporated into a middleware framework based on the CORBA Component Model (CCM), which allows coarse grain parallelization and distributed computing, among other things. Specifications have been written for each of the transport solutions, which provide descriptions of the solution techniques at a level of detail not previously available. The reengineered versions are written primarily in Java, while the original legacy versions were written in FORTRAN 77. Although there have been concerns that Java is slower than FORTRAN, these concerns have been allayed, since the actual difference in processing speed has been found to be within a factor of 2, which can be offset by the parallel processing capability. Certain legacy models are incorporated directly using wrappers rather than being rewritten. The new software is embedded in the Enterprise Java CORBA Component Model, an implementation of the CCM that was developed at CPI. The new modeling software offers improved maintainability, more extensive code reuse where similar algorithms are used in different models, and greater flexibility for studying terrestrial and planetary atmospheres. It is being added to a more general suite of open source software called Aeronomy Modeling Phenomenology Tool (APMT), which also includes a photochemistry model developed under a previous AISR project.

Climatological Database of Auroral Images for Solar Cycle 23: An Online Synoptic Search and Metadata

PI: Glynn A. Germany, University of Alabama, Huntsville

We propose to provide an online search and visualization tool of auroral and geophysical metadata covering the ascending phase of the current solar cycle. This tool will address a demonstrated problem in using auroral image collections, namely the extreme difficulty of performing non-event studies and surveys of large (~4M) collections of auroral images. The tool will be based on the collection of POLAR UVI images available from March 1996 to present, but will allow ingestion of metadata from the other POLAR imagers (VIS and PIXIE) as well as other imaging missions such as the imager suite on IMAGE. The proposed work directly addresses a major problem associated with all the POLAR imagers (and spaced-based imagers in general), namely that data is typically available only on an event-driven basis. There is currently no tool available for synoptic searches of auroral features or for visualization of derived parameters such as boundaries, energy input, or presence of given auroral morphologies. We propose to provide that tool. The methodology is feasible and appropriate to the proposal goals, and has been explored in sufficient detail to insure its success. The methodology builds on previous success with image classification techniques and development of user interfaces to mission databases essentially identical to that proposed here. In summary, this work builds on previous experience with analysis of auroral image morphologies and with online access to scientific databases, uses a team intimately familiar with the data sets to be enhanced, addresses a significant problem within the magnetospheric community that is directly relevant to the AISR goals and areas of interest, promises significant results early in the funding period, and has a high probability of successfully achieving the proposed goals.

Automated Event Detection and Classification in High Volume Space Physics Time Series Data using Machine Learning

PI: Simon Perkins, Los Alamos National Laboratory

Modern satellite-based space physics instruments generate vast streams of temporal multi-band data. A trained expert can often examine a suitably displayed portion of the data stream and identify such events by hand, but for high-volume data streams this approach is entirely impractical. Writing robust feature detectors is very time-consuming and requires a high degree of programming skill and familiarity with pattern recognition techniques -- certainly not skills that every scientist has. We propose to investigate the use of supervised machine learning techniques to make this job much easier. We intend to develop a general purpose user-friendly software tool for automatically generating event classifiers that operate on multi-band time-series data using advanced machine learning techniques. We intend to look at using the relatively recently developed technique of Support Vector Machines. SVMs have a strong mathematical foundation and have been empirically shown to be one of the most powerful powerful and robust machine learning technologies available today. We will also consider the use of Boosting and Gaussian Processes, two other recently developed and mathematically justifiable machine learning paradigms. Using whichever technique works best, we will develop a general software tool that can be applied to a variety of different space physics data streams. The primary target of our initial research will be magnetospheric data collected by satellites such as the Los Alamos geosynchronous satellites (Magnetospheric Plasma Analyzer, MPA; energetic particles, SOPA and ESP), The GPS energetic particle detectors, with further application to other data sets existing in-house (POLAR CEPPAD/CAMMICE/HYDRA) as well as upcoming data from the recently launched Cluster II sets of satellites.

Development, Visualization, Retrieval, and Analysis of Multispacecraft Data Sets

PI: D. Aaron Roberts, Goddard Space Flight Center

We propose to produce and disseminate new methods for the retrieval, visualization, and analysis of multi-spacecraft data. Data set browsing will initially be made possible through the use of uniform, averaged data sets, but the software will be compatible with multi-platform, distributed data retrieval. The work builds on previous visualization and data analysis tools from AISRP-supported and other work, and will be applicable to nearly all NASA space physics missions, including ISTP, Magnetospheric Constellation and related missions, and the "COHOWeb" missions to the heliosphere. The results will be incorporated into the National Space Science Data Center (NSSDC) data dissemination tools, allowing scientists and the public to visualize spacecraft data in the 3-D context that it was measured (tied to the spacecraft trajectories) and to download data intervals of interest in directly useable form (interpolated, despiked, etc.). The software will be developed using common industry tools (e.g., Java3D, OpenGL, C, C++, Fortran, and X) and will be linked to very extensively used commercial packages such as IDL. Server-based versions of the software will be used for browsing and data set preparation, and stand-alone routines will be available to analyze the data after download. A variety of file formats will be available for downloaded data, and the software will be as platform-independent as possible. To add to the utility of the project, we will develop a parallel processing version of the retrieval software, so that more and larger data requests can be handled efficiently.

Statistical Object Identification, Tracking and Analysis

PI: Michael J. Turmon, Jet Propulsion Laboratory

Scientists are overwhelmed by the vast amounts of data available from instruments, experiments, and simulations; all information technology trends point to a continuing increase in the size of these datasets. Unfortunately, the ability to collect, store, and manipulate ever-greater volumes of data will not translate into greater scientific understanding without development of suitable software. One way to conceptualize the process of discovering patterns in these large spatiotemporal datasets is to identify objects, which we take to be spatially coherent regions of interest that evolve through time. This provides one way for scientists to rise above a pixel-level understanding of their data to a more natural and physically relevant description in terms of salient features. To analyze objects, they must be identified in individual images, tracked through image sequences, and their temporal behaviors modeled quantitatively. We may solve a large class of object-identification problems with modern, general-purpose spatial statistics methods which employ unsupervised learning or scientist-provided labelings to define local assessments of activity. These pixel-level cues are progressively linked into a coherent, object-level scene description. Similarly general statistical time series models are used for object tracking and analysis. Expressing the models in a neutral, mathematical/statistical language allows repeatability, model refinement over time, and model exchange among investigators. Our use of open standards for structured data (e.g., object models and datasets in XML with suitable schemata) ensures portability across applications and future extensibility. The system is field-tested on a multiwavelength series of solar images in conjunction with solar physicists. Strong collaborations with domain scientists ensure successful deployment into other Sun-Earth connection missions, and OSS application areas like planetary science and geophysics.

Bibliography

Statistical Pattern Recognition for Labeling Solar Active Regions: Application to SOHO/MDI Imager Turmon; Pap; Mukhtar; ApJ, (2002) 568 pp. 396

Package for Interactive Analysis of Line Emission (Analysis of UV--X-Ray High-Resolution Emission Spectra)

PI: Vinay L. Kashyap, Harvard/Smithsonian Center for Astrophysics

PINTofALE is an IDL based package to analyze high-resolution grating spectra. The first version was made available to the public on 3 February 2001. Since then we have carried out numerous changes and subsidiary releases. The current release is version 2.0 (released 6 Apr 2004), and we are preparing to release v2.1 within the next month. The changes include bug fixes, upgrades to handle higher versions of IDL and the CHIANTI database, enhancements in user-friendliness, handling of instrument response matrices, and the release of a Markov Chain Monte Carlo based DEM fitting routines. A detailed description of the package, together with fairly detailed documentation, example walks-throughs, and downloadable tar files, are available on-line from <http://hea-www.harvard.edu/PINTofALE/>. The website also lists papers that have used PINTofALE in their analysis.

Tomographic analysis and visualization of remotely-sensed heliospheric data

PI: Bernard V. Jackson, University of California San Diego

The proposed project consists of two intimately related tasks. The first is the development of a generic tomographic deconvolution tool, capable of reconstructing the global, 3D density and velocity structure of the solar wind in the inner heliosphere using remote sensing data. The technique takes a novel approach by including outward motion of structures in the solar wind to obtain the different perspectives required for a tomographic analysis. Observations from a single vantage point are needed only, making the technique applicable to e.g. interplanetary scintillation (IPS) observations, Thomson scattering data from the HELIOS photometers, coronagraphs (e.g. SOHO/LASCO), and future all-sky imagers such as the Solar Mass Ejection Imager (SMEI) and imagers on missions such as STEREO, Solar Probe and the Solar Polar Sail Mission. Proper understanding of tomographic results requires advanced visualization tools designed to deal with a 3D plasma environment in a heliospheric context. The second task is the development of tools for displaying and manipulating the 3D tomography results. This will target a system capable of simultaneously displaying density, velocity and magnetic field, and will allow the user to interactively explore the heliospheric model from any desired perspective. We will also experiment with advanced techniques, such as stereoscopy, drawing on the experience and hardware at the VisLab of the San Diego Supercomputer Center (SDSC). The proposed tools have the potential to become crucially important to researchers active in heliospheric remote sensing. The CASS scientists on this proposal have worked with remote sensing data for many years, including HELIOS photometer and IPS observations, and, more recently, the SOHO/LASCO coronagraph data. They are involved in the construction of SMEI, and have contributed to the design of prospective all-sky imagers for Solar Probe and the Solar Polar Sail Mission as co-investigators in NASA-funded concept studies.